

Use of NGS for susceptibility testing



**Henrik Hasman
Senior Scientist
Statens Serum Institut
Denmark**



Overview

- 1. AMR and NGS crash course**
- 2. NGS tools for AMR detection**
- 3. Examples and challenges in relation to NGS/AMR**



Resistance to Antimicrobial Agents

From genotype to phenotype

- **Resistance mechanisms**

- Some bacteria are naturally (intrinsic) resistant

- Lack/inaccessibility of target for antimicrobial agent
- Chromosomal resistance gene
- Multi-drug pumps
- Biofilm production
- And much more.....

- Acquired resistance

- Mutations of chromosomal genes/targets
- Acquisition of resistance genes e.g. on plasmids or transposons

Additional challenge: Genes may not be well expressed.



Known gene/mutation → predicted phenotype

Known knowns, Known unknowns and unknown unknowns*

“Known knowns”

- published and well-characterized AMR gene sequences or mutations

CTX-M-15 → beta-lactamase with extended spectrum → CTX^R (CAZ^R, MERO^S..)

“Known unknowns”

- gene variants of known AMR gene sequences

TEM-345 → beta-lactamase unknown spectrum → CTX[?] (CAZ[?], MERO[?]..)

“Unknown unknowns”

- new threats emerging e.g. from the environment

Mcr-1 → Phospho-ethanolamine transferase → Reduced susceptibility to colistin

* Credit for the “known knowns,” “known unknowns,” and “unknown unknowns” perspective on the AMR crisis is given to Dr. Gerard D. Wright of McMaster University’s Michael G. De Groote Institute for Infectious Disease Research. The original quote is attributed to former *United States Secretary of Defense* Donald Rumsfeld.



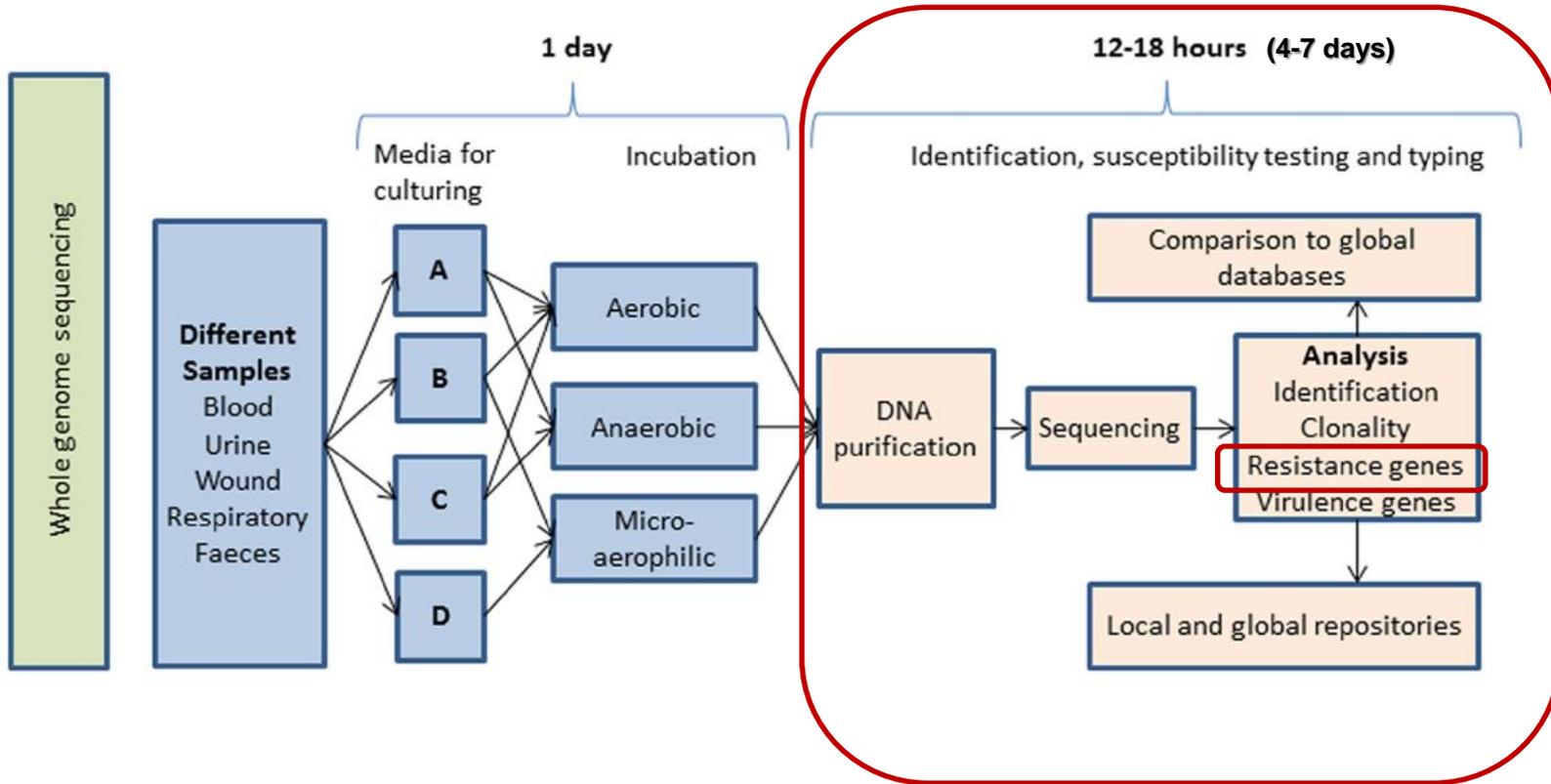
Define why you want NGS!

What is the purpose for using NGS in relation to AMR?

- To predict phenotypes based on genotypes?
- To support bacterial typing data?
- To perform AMR surveillance (local, national or global)?
- To publish...?



Workflow with NGS at the clinical laboratory



How do I do that in my laboratory?



Become a bioinformatician!

```
Chr1 3395973 rs143478237 ACC A 83 PASS DB:DP=9;HP=5;NF=3;NFS=3;NR=0;NRS=0;VCQ=INTRONIC;VCQNC=INTRONIC;VGN=ARHGEF16 GT:GQ 1/1:9
Chr1 3545175 rs147637374 GTCTGGGAGCTCCTCCCC G 141 PASS DB:DP=55;HP=2;NF=2;NFS=5;NR=5;NRS=13;VCQ=SPLICE_SITE:3PRIME_UTR;VCQNC=UPSTREAM;VGN=TPRGL1 GT:GQ 0/1:99
Ingos-MacBook-Pro:NLIS3_4 ih$ sed 50q 2.vcf
##fileformat=VCFv4.1
##ALT=<ID=DEL,Description="Deletion">
##FILTER=<ID=fr0,Description="Non-ref allele is not covered by at least one read on both strands">
##FILTER=<ID=hp10,Description="Reference homopolymer length was longer than 10">
##FILTER=<ID=q2F,Description="Reads covering non-ref variant site on forward strand">
##FILTER=<ID=q2R,Description="Reads covering non-ref variant site on reverse strand">
##FILTER=<ID=w,Description="Reads covering non-ref variant site with higher likelihood">
##FORMAT=<ID=GC,Description="GC content">
##FORMAT=<ID=GT,Description="Genotype">
##INFO=<ID=DB,Number=1,Type=String,Description="Database membership">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Read depth">
##INFO=<ID=HP,Number=1,Type=Integer,Description="Number of reads in haplotype window">
##INFO=<ID=NF,Number=1,Type=Integer,Description="Number of reference homopolymer tract length">
##INFO=<ID=NF,Number=1,Type=Integer,Description="Number of reads covering non-ref variant on forward strand">
##INFO=<ID=NFS,Number=1,Type=Integer,Description="Number of reads covering non-ref variant site on forward strand">
##INFO=<ID=NR,Number=1,Type=Integer,Description="Number of reads covering non-ref variant on reverse strand">
##INFO=<ID=NRS,Number=1,Type=Integer,Description="Number of reads covering non-ref variant site on reverse strand">
##VariantAnnotation
##INFO=<ID=VEP,Number=1,Type=String,Description="Ensembl VEP (from ensembl VEP) - non coding genes (if available and only transcripts without a translation are included). Highest splice consequence (if applicable), followed by most severe consequence in the VCQ field (if applicable). Note one is selected.">
##INFO=<ID=VPI,Number=1,Type=String,Description="VPI (from ensembl VEP) associated with the most severe consequence in the VCQ field (if applicable). Note one is selected.">
##INFO=<ID=Condel,Number=1,Type=String,Description="Condel (from ensembl VEP) for the most severe consequence in the VCQ field (if applicable).">
##INFO=<ID=VGN,Number=1,Type=String,Description="VGN (from ensembl VEP) associated with the most severe consequence in the VCQ field.">
##INFO=<ID=VAA,Number=1,Type=String,Description="VAA (from ensembl VEP) associated with the most severe consequence in the VCQ field (if applicable).">
##INFO=<ID=PolyPhen,Number=1,Type=String,Description="PolyPhen (from ensembl VEP) for the most severe consequence in the VCQ field (if applicable).">
##INFO=<ID=VP,Number=1,Type=String,Description="VP (from ensembl VEP) - coding genes (if available and only transcripts with a translation are included). Highest splice consequence : (if applicable), followed by most severe consequence in the CQ field (if applicable).">
##INFO=<ID=SIFT,Number=1,Type=String,Description="SIFT (from ensembl VEP) for the most severe consequence in the CQ field (if applicable).">
##INFO=<ID=VPP,Number=1,Type=String,Description="VPP (from ensembl VEP) associated with the most severe consequence in the VCQ field (if applicable).">
#CHROM POS ID REF ALT QUAL INFO
Chr1 866511 rs146519568 C CCCC NRS=1;NR=1;NRS=7;VCQ=INTRONIC;VGN=SAMD11 GT:GQ 0/1:13
Chr1 874950 rs149166309 T TCCC NRS=3;NF=0;NFS=0;NR=4;NRS=4;VCQ=INTRONIC;VCQNC=UPSTREAM;VGN=SAMD11 GT:GQ 1/1:21
Chr1 907763 . G GTC 1 INTRONIC;VCQ=UPSTREAM;VGN=PLEKH11;VPI=ENSP0000368719;VPP=373 GT:GQ 0/1:1
Chr1 948846 rs3841266 T TA NRS=23;NR=1;NRS=1;VCQ=SPRIME_UTR;VCQNC=UPSTREAM;VGN=ISG15 GT:GQ 1/1:75
Chr1 957967 rs141489152 T TTGTA 35 DB:DP=20;HP=2;NF=1;NFS=1;NR=1;NRS=2;VCQ=INTRONIC;VGN=AGRN GT:GQ 1/1:28
Chr1 978549 rs56001364 TG T 12;NRS=22;NR=2;NRS=2;VCQ=INTRONIC;VCQNC=INTRONIC;VGN=AGRN GT:GQ 1/1:11
Chr1 978603 rs146114193 CCT C VCQ=INTRONIC;VCQNC=UPSTREAM;VGN=AGRN GT:GQ 0/1:91
Chr1 984171 rs140904842 CAG C VCQ=INTRONIC;VCQNC=WITHIN_NON_CODING_GENE;VGN=AGRN GT:GQ 0/1:7
Chr1 999841 rs144946318 ATG A VC=INTRONIC GT:GQ 1/1:5
Chr1 1158534 rs59317408 . VCQ=INTRONIC;VCQNC=UPSTREAM;VGN=SDF4 GT:GQ 0/1:99
Chr1 1158562 rs139833693 AAC A VCQ=INTRONIC;VCQNC=UPSTREAM;VGN=SDF4 GT:GQ 0/1:99
Chr1 1276973 rs145370195 G GACAC 473 VCQ=INTRONIC;VCQNC=DOWNSTREAM;VGN=DVL1 GT:GQ 1/1:6
Chr1 1289367 rs140777846 CTG C VCQ=INTRONIC;VCQNC=INTRONIC;VGN=MXR8 GT:GQ 1/1:72
Chr1 1323143 rs147796530 CCT C VCQ=INTRONIC;VCQNC=INTRONIC;VGN=CCNL2 GT:GQ 1/1:30
Chr1 1325493 . CA C 23 hp10 VCQNC=INTRONIC;VGN=CCNL2 GT:GQ 1/1:16
Chr1 1588744 rs79724854 AGCG A 133 VCQ=DOWNSTREAM;VCQNC=INTRONIC;VGN=SLC35E2B GT:GQ 0/1:99
Chr1 1633003 . C CCG 154 PASS VCQNC=INTRONIC;VGN=CDK11A GT:GQ 1/1:15
Chr1 1647649 rs70931719 C CT 410 VCQ=INTRONIC;VCQNC=INTRONIC;VGN=CDK11A GT:GQ 1/1:29
Chr1 1647893 rs144636354 C CTTCTT 1134 PASS DB:DP=236;HP=3;NF=1;NFS=3;NR=30;NRS=33;VAA=R/KER;VCQ=NON_SYNONYMOUS_CODING;VCQNC=WITHIN_NON_CODING_GENE;VGN=CDK11A;VPI=ENSP00000422149;VPP=1
27 GT:GQ 0/1:99
Chr1 1647968 rs146207119 C CAT 7 q20 DB:DP=115;HP=1;NF=3;NFS=4;NR=5;NRS=32;VCQ=INTRONIC;VCQNC=INTRONIC;VGN=CDK11A GT:GQ 0/1:7
Chr1 1650639 . CCA C 35 PASS DP=124;HP=2;NF=2;NFS=15;NR=0;NRS=0;VCQ=INTRONIC;VCQNC=INTRONIC;VGN=CDK11A GT:GQ 0/1:35
Chr1 1653332 rs34272957 GT G 5 q20 DB:DP=8;HP=2;NF=0;NFS=0;NR=1;NRS=1;VCQ=INTRONIC;VCQNC=INTRONIC;VGN=CDK11A GT:GQ 1/1:4
Ingos-MacBook-Pro:NLIS3_4 ih$
```

UNIX

Python

Perl

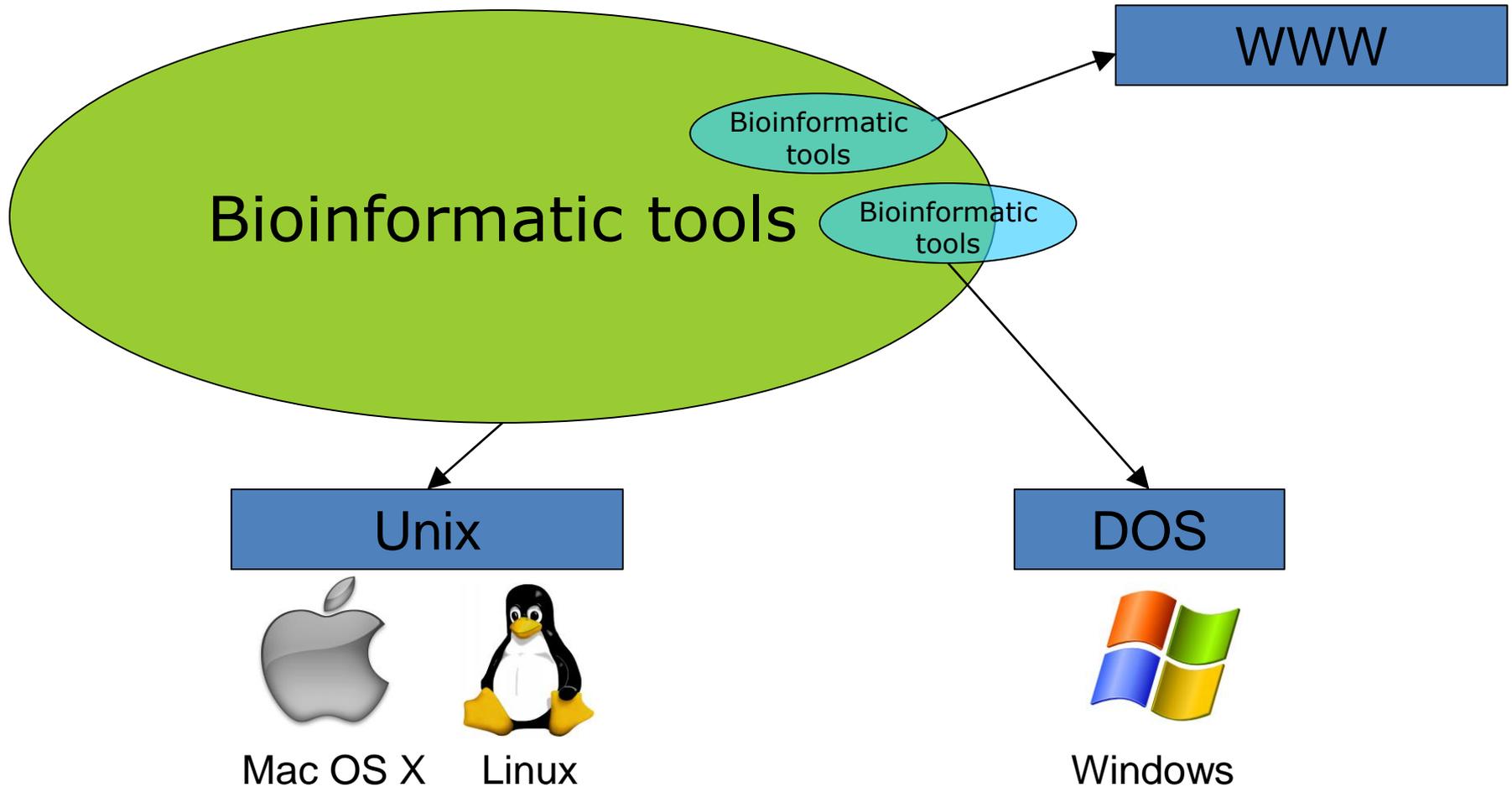
JAVA

JAVAScript



Introduction & Background

NGS data analysis – Bioinformatics



Mandatory readings

Edwards and Holt *Microbial Informatics and Experimentation* 2013, 3:2
<http://www.microbialinformaticsj.com/content/3/1/2>



REVIEW

Open Access

Beginner's guide to comparative bacterial genome analysis using next-generation sequence data

David J Edwards^{1,2} and Kathryn E Holt^{1*}

phg foundation
making science work for health

Pathogen Genomics Into Practice

July 2015



Take a course...

coursera

Catalog

Search catalog



For Enterprise

Log In

Sign Up

Home > Life Sciences > Clinical Science

Whole genome sequencing of bacterial genomes – tools and applications

Overview

Syllabus

FAQs

Creators

Ratings and Reviews

Enroll

Started Apr 30

[Apply for Financial Aid](#)

About this course: This course will cover the topic of Whole genome sequencing (WGS) of bacterial genomes which is becoming more and more relevant for the medical sector. WGS technology and applications are high on international political agenda, as the classical methods are being replaced by WGS technology and therefore bioinformatic tools are extremely important for allowing the people working in

[More](#)

Who is this class for: This course is for you if you are interested in getting to know more about Whole genome sequencing applied to bacterial characterization and surveillance. We aim at having a broad scope and international reach in different sectors. So this course is for you whether you are an undergraduate or graduate student, a researcher, medical or veterinary related professional, technical staff or simply interested in the subject!

Created by: Technical University of Denmark (DTU)



Take a course...



ESCMID

MANAGING INFECTIONS
PROMOTING SCIENCE

ESCMID Postgraduate
Technical Workshop

**Capacity-Building
Workshop: Whole-Genome
Sequencing for Clinical
Microbiology and Hospital
Epidemiology**

**Freiburg, Germany
10 – 13 October 2017**



GoSeqIt

Sequence based microbial diagnostics

HOME

WORKSHOPS ▾

WEB-SERVICES

SCIENTIFIC FOUNDATION



Workshops

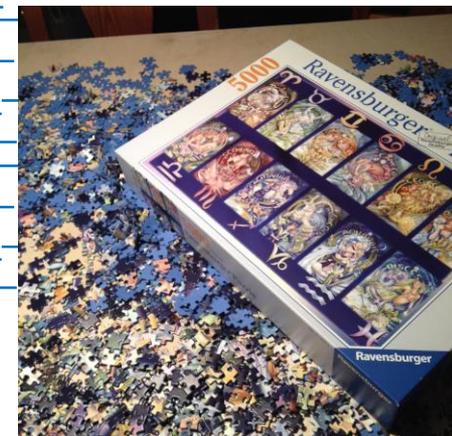
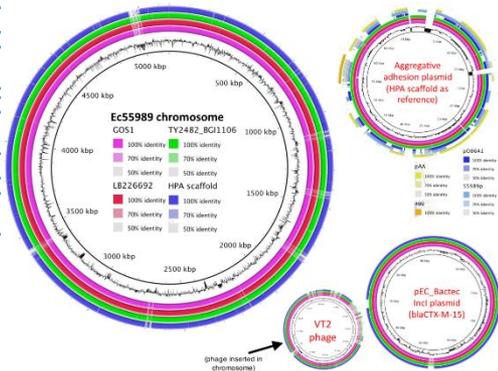
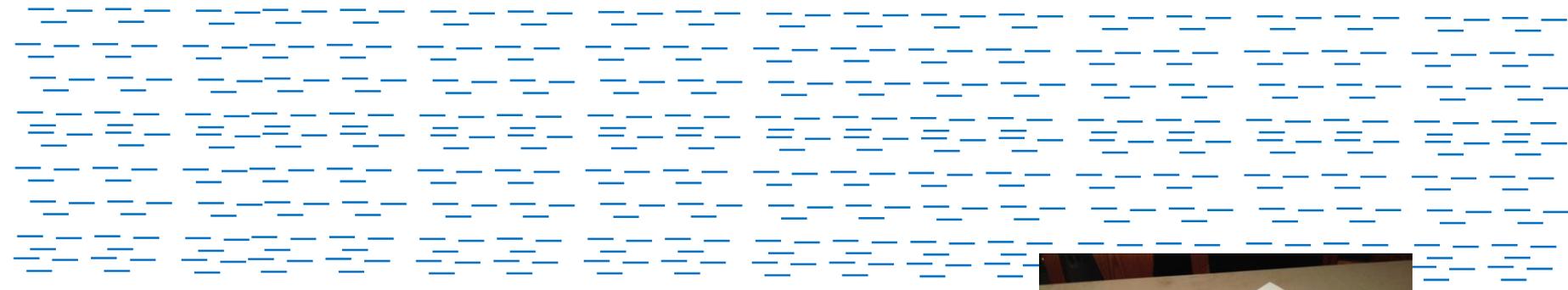
Sign up for workshop Sep. 2018



NGS Crash course

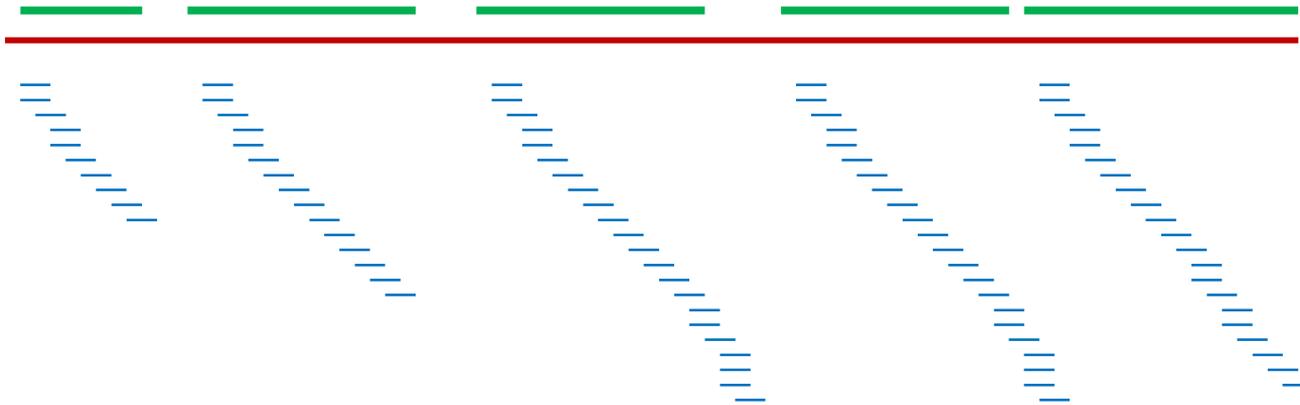
in relation to AMR

Huge numbers of small fragments (100-500 bp)



Reference vs. *de novo* assembly

Reference assembly (mapping)



Known genome



De novo assembly



Draft genome



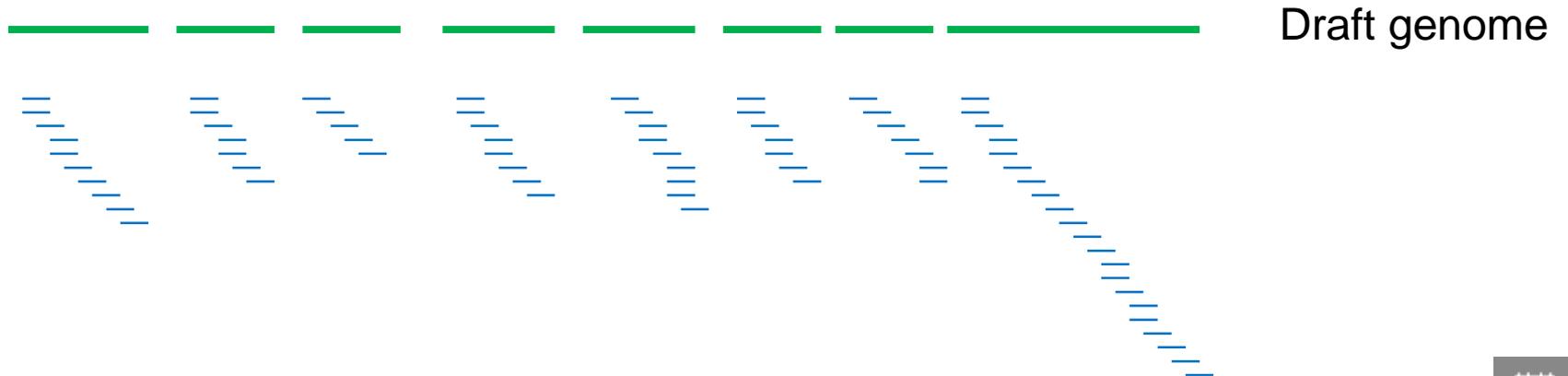
De novo assembly



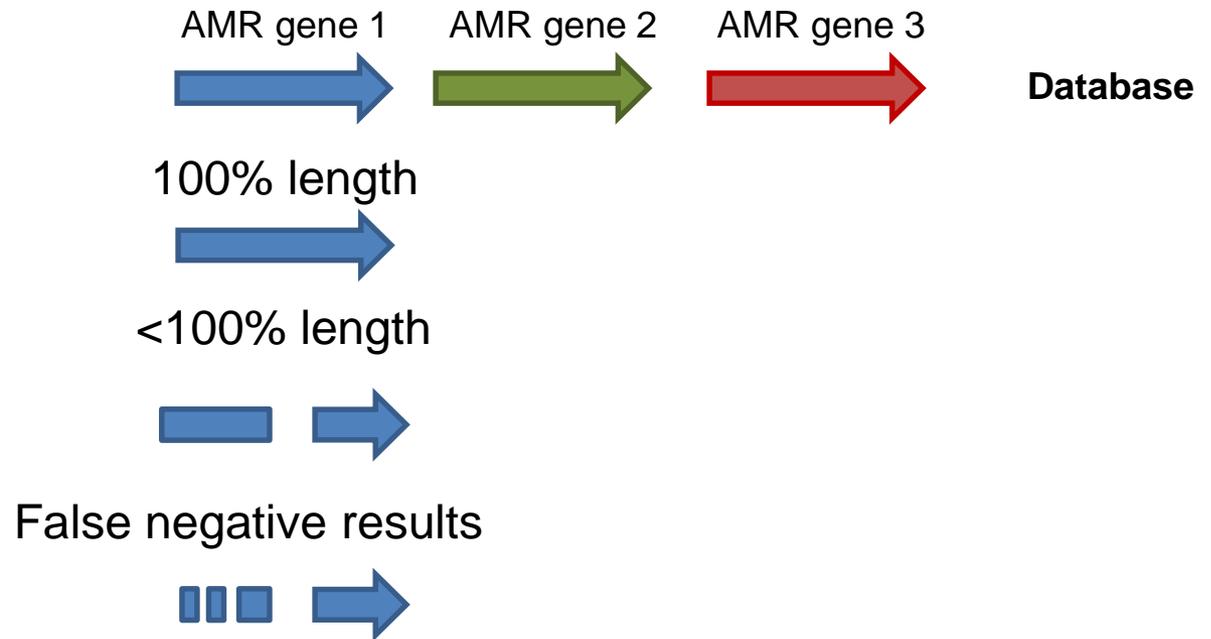
Gene detection: BLAST alignment towards curated gene database

De novo Mutations

Sequence scanning towards curated mutation database



Pitfalls for de novo detection of AMR

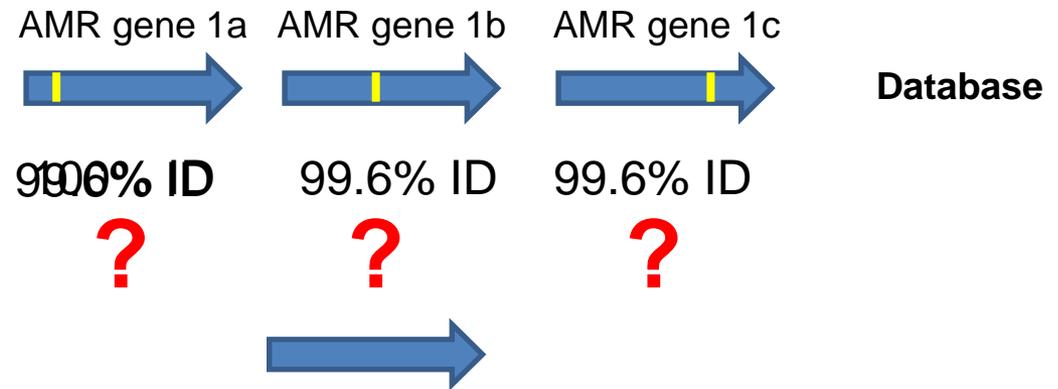


De novo assembly requires high depth of coverage to ensure high sensitivity

...and is therefore less sensitive to traces of contaminating DNA



Pitfalls for de novo detection of AMR

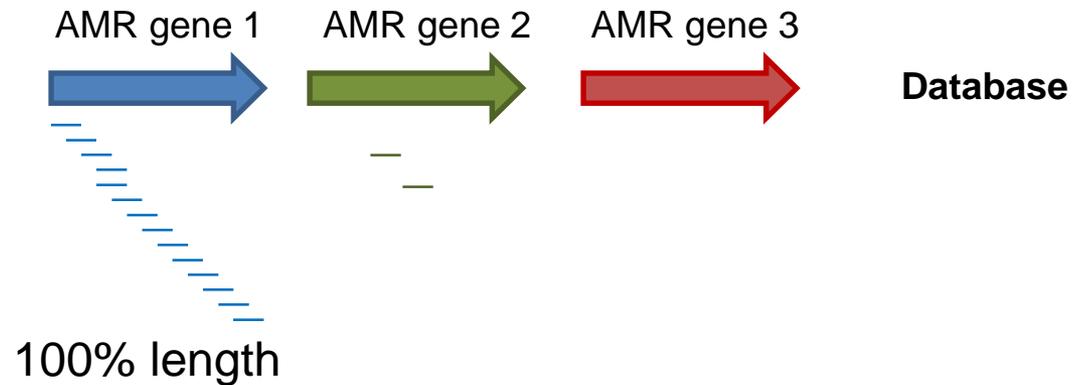


Can we be sure that a gene <100% identical to the published genes has the same phenotype?

≠2 nt $\left\{ \begin{array}{l} \text{ACC(6')-Ib} \rightarrow \text{aminoglycoside resistance} \\ \text{ACC(6')-Ib-cr} \rightarrow \text{aminoglycoside + low-level fluoroquinolone resistance} \end{array} \right.$



Pitfalls for detection of AMR by mapping



Mapping increases sensitivity for detecting AMR genes

...including traces of contaminating DNA

Important note: *Make sure your NGS data has sufficient quantity and quality before you start working with them!*



EXPERT
MANAGEMENT TRAINING
CERTIFICATE OF COMPLETION

THIS IS TO CERTIFY THAT:

YOUR NAME HERE

HAS SUCCESSFULLY COMPLETED:

NGS in AMR detection – a crash course

May 8. 2018

Date of Issue



M. Kakead
Course Director

www.expertmanagementtraining.com



How to choose the appropriate tool?

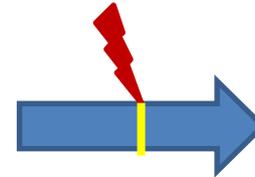
1. Platform

- UNIX?
- Windows?
- MAC?
- Web-based?



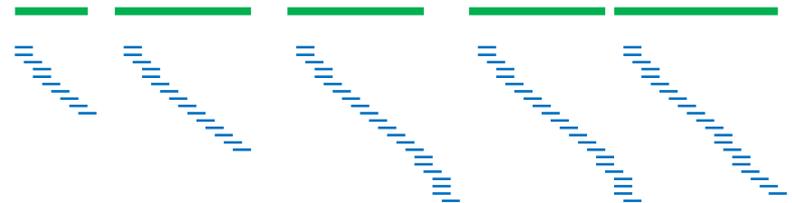
2. Database

- Curated (and updated?) or home-brew database?
- Genes (all/transerable)?
- Mutations?
- Species specific?



3. Detection method

- Mapping versus *de novo*?
- *Specific or general tools*?



Important note: *Validate your tools (pipelines) with a well-characterized data set before you implement it in your daily work!*



Mandatory readings

Clinical Microbiology and Infection 23 (2017) 2–22

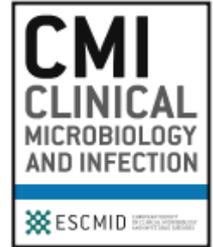


ELSEVIER

Contents lists available at [ScienceDirect](#)

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com



Review

The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee

M.J. Ellington^{1,†}, O. Ekelund^{2,†}, F.M. Aarestrup³, R. Canton⁴, M. Doumith¹, C. Giske⁵, H. Grundman⁶, H. Hasman⁷, M.T.G. Holden⁸, K.L. Hopkins¹, J. Iredell⁹, G. Kahlmeter², C.U. Köser¹⁰, A. MacGowan¹¹, D. Mevius^{12,13}, M. Mulvey¹⁴, T. Naas¹⁵, T. Peto¹⁶, J.-M. Rolain¹⁷, Ø. Samuelsen¹⁸, N. Woodford^{1,*}



Mandatory readings



Molecular Methods for Detection of Antimicrobial Resistance

MUNA F. ANJUM,¹ EA ZANKARI,² and HENRIK HASMAN^{2,3}

¹Department of Bacteriology, Animal and Plant Health Agency, Surrey, United Kingdom;

²National Food Institute, Technical University of Denmark, Lyngby, Denmark; ³Reference Laboratory for Antimicrobial Resistance and Staphylococci, Staten Serum Institut, Copenhagen, Denmark



Mandatory readings

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *Antimicrobial Therapeutics Reviews*

Antimicrobial resistance surveillance in the genomic age

Andrew G. McArthur and Kara K. Tsang

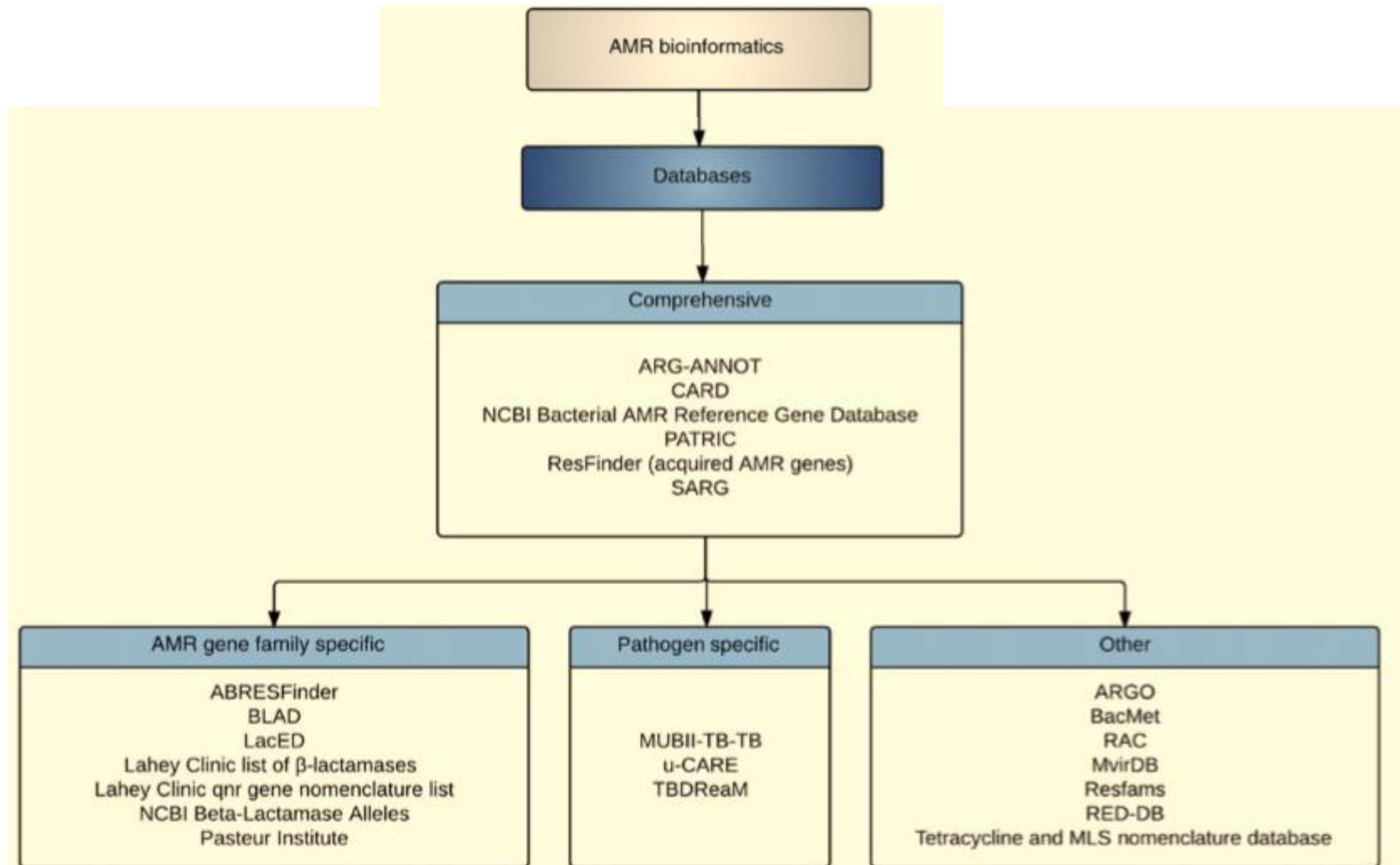
M. G. DeGrootte Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGrootte School of Medicine, McMaster University, Hamilton, Canada

Address for correspondence: Andrew G. McArthur, M. G. DeGrootte Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGrootte School of Medicine, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. mcarthua@mcmaster.ca

Important note: *Remember to also read carefully the specific literature describing any tool you wish to use for detection of AMR*



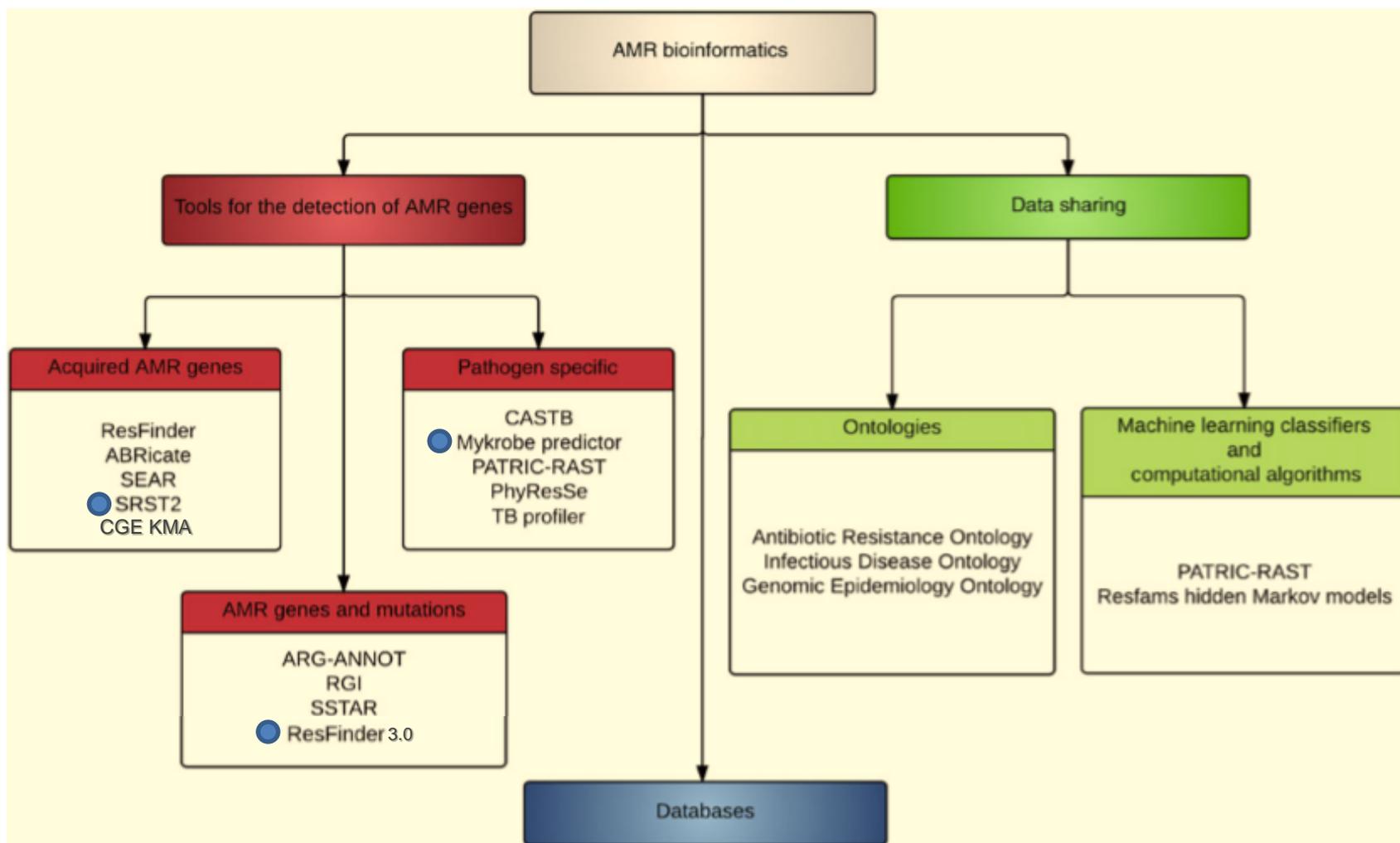
Overview - Databases



Overview - Tools

McArthur & Tsang

AMR surveillance in the genomic age



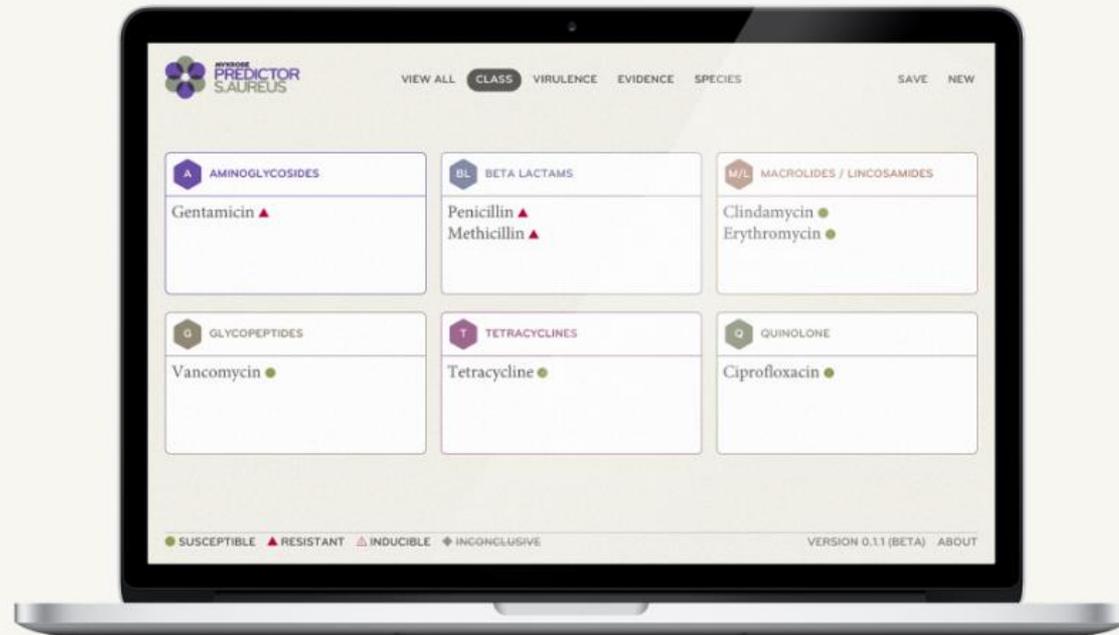
Pathogen specific tools

MYKROBE

HOME PREDICTOR ABOUT CONTACT

PREDICTOR

The Mykrobe predictor is designed for use by microbiologists and doctors, providing information needed in order to choose the best treatment. It analyses the whole genome of a bacterial sample, all within a couple of minutes, and predicts which drugs the infection is resistant to. No expertise is needed to run or interpret it, and it works on a standard desktop or laptop.



MRSA

TB



TECHNICAL SPECS

Runs on Windows, Mac OS X and Linux, requires ~300Mb of RAM, takes between 40 seconds and 3 minutes to run.

Sequencing: Supports Illumina sequencing data as



CARD website

The Comprehensive Antibiotic Resistance Database

A bioinformatic database of resistance genes, their products and associated phenotypes.

3821 Ontology Terms, 2388 Reference Sequences, 1046 SNPs, 2394 Publications, 2420 AMR Detection Models

Browse

The CARD is a rigorously curated collection of known resistance determinants and associated antibiotics, organized by the Antibiotic Resistance Ontology (ARO) and AMR gene detection models.

Analyze

The CARD includes tools for analysis of molecular sequences, including BLAST and the Resistance Gene Identifier (RGI) software for prediction of resistome based on homology and SNP models.

Download

CARD data can be downloaded in a number of formats. RGI software is available as a command-line tool. View changes in CARD curation. CARD is updated monthly.

Prevalence

A new computer-generated data set on the prevalence of AMR genes and variants among the sequenced genomes, plasmids, and whole-genome shotgun assemblies available at NCBI for clinically important pathogens.

Timeline

 Joseph Nesme Retweeted 

 **CARD Developers**
@arpcard

[#card_release](#) CARD 1.2.1: MCR-1.2, beta-lactamases, APHS + Prevalence 1.0.2 data set based on August 2017 genomes card.mcmaster.ca/home

  Oct 13, 2017

<https://card.mcmaster.ca/>



CARD website

Analyze

Single gene upload

BLAST

Perform standard BLAST searches against the CARD reference sequences. Results are annotated with extra information from the CARD.

RGI

Use the Resistance Gene Identifier to predict resistome(s) from protein or nucleotide data based on homology and SNP models. Now supporting bulk analyses.

[Visualize external RGI json results](#)

Draft genome upload

RGI Resistance Gene Identifier

RGI Bulk Analysis: Do you want to use RGI to analyze a large number of genomes? It is now available as a downloadable command-line tool in the [Download section of the CARD website](#).

RGI 3.2.1: Open Reading Frame (ORF) prediction using [Prodigal](#), homolog detection using [Diamond](#), and Strict significance based on CARD curated bitscore cut-offs.

The CARD includes a curated collection of antimicrobial resistance gene and mutation sequences, bioinformatics models for their detection, and software in the form of the Resistance Gene Identifier (RGI) for their detection in genome or protein sequences. It includes a common framework for the sharing of antimicrobial resistance data in the form of the novel Antibiotic Resistance Ontology (ARO), allowing the RGI to catalog resistome predictions by both drug class and resistance mechanism.

The RGI provides a preliminary annotation of your DNA or protein sequence(s), based upon the data available in CARD. The RGI currently only analyzes protein sequences - if genome sequences or assembly contigs are submitted, the RGI first predicts open reading frames using the [Prodigal](#) software (ignoring those less than 30 bp) and analyzes the predicted protein sequences. The RGI currently supports two detection model types - *protein homolog models* (use of sequence similarity cut-offs to detection functional homologs of AMR genes) and *protein variant models* (for accurate differentiation between susceptible intrinsic genes and intrinsic genes that have acquired mutations conferring AMR, based on curated SNP matrices).

Disclaimer: As the CARD curation evolves, the results of the RGI evolve. RGI targets, reference sequences, and significance cut-offs are under constant curation.

The RGI analyzes sequences under three paradigms - Perfect, Strict, and Loose (a.k.a. Discovery). The Perfect algorithm is most often applied to clinical surveillance as it detects perfect matches to the curated reference sequences and mutations in the CARD. In contrast, the Strict algorithm detects previously unknown variants of known AMR genes, including secondary screen for key mutations, using detection models with curated similarity cut-offs to ensure the detected variant is likely a functional AMR gene. **The Loose algorithm works outside of the detection model cut-offs to provide detection of new, emergent threats and more distant homologs of AMR genes, but will also catalog homologous sequences and spurious partial hits that may not have a role in AMR.** Combined with phenotypic screening, the Loose algorithm allows researchers to hone in on new AMR genes.

New: The online RGI now caches results for each session for 7 days, allowing you to work on multiple data sets using the web interface.

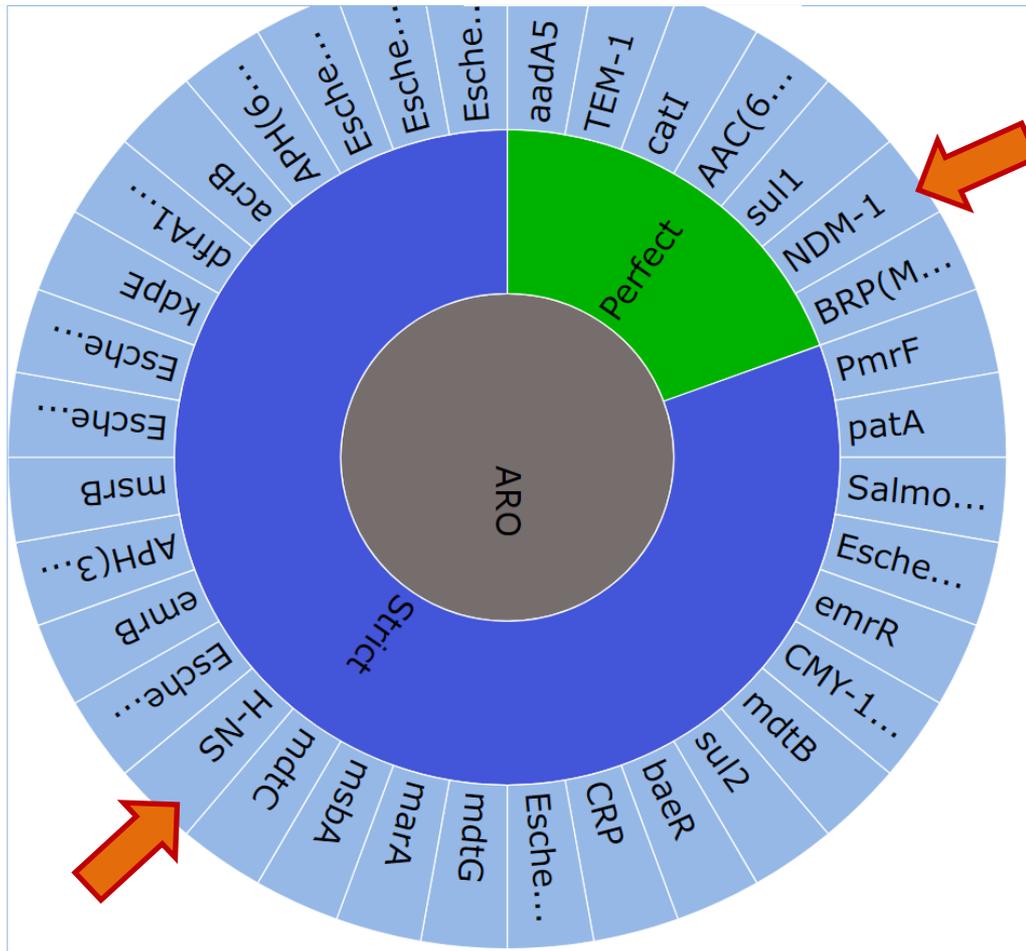


CARD website

Resistance Gene Identifier (RGI) Results

Loose hits of e-10 or better can be visualized in the "View by Functional Categories" display. Loose hits are not shown in the "View by Gene" display.

Filename	Date (UTC)	Criteria	Visualize	All Results (JSON)	Tab-delimited Summary	FASTA Format	Heat Map PDF
Strain1_contigs.fasta	January 12, 2018 09:24:22	Perfect, Strict	View by Gene View by Functional Categories	Download	Download	Download	Download



CARD website

[Perfect]

Name	Value
ARO Accession	3000589
ARO Category	determinant of beta-lactam resistance,antibiotic inactivation enzyme
bit score	544.3
E-value	2.9e-156
Percentage Identity (of matching region)	100%
Percentage Identity (of matching region)	100%
Model type	protein homolog model
Model Cut-off	500
Source	Strain1_NODE_33_length_11874_bpcov_19.9_2
label	Strain1_NODE_33_length_11874_bpcov_19.9_2#626#1438#-1#ID=33_2;partial=00;start_type=ATG;rbs_motif=AGGA;rbs_spacer=5-10bp;gc_cont=0.614
id	gn BL_ORD_ID 914 hsp_num:0
Start	626
End	1438
Strand	-

Predicted Sequence

```
MELPNIHHPVAKLSTALAAALMSGCHPGEIRPTIGQQMETGDQRFGLVFRQLAPNVWQ
HTSYLDMPGFGAVASNLIVRDGGRVLVVDTAWTDQTAQILNWIKQEINLPVALAVVTH
AHQDKMGGMDALHAAGIATYANALSNQLAPQEGHVAQAQSLTFAANGWEPATAPNFGPL
KVFPYGPGHSDNITVGDGTDIAFGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAF
PKASHIVMSHSAPDSRAAITHTARMADKLR
```

CARD Sequence

```
MELPNIHHPVAKLSTALAAALMSGCHPGEIRPTIGQQMETGDQRFGLVFRQLAPNVWQ
HTSYLDMPGFGAVASNLIVRDGGRVLVVDTAWTDQTAQILNWIKQEINLPVALAVVTH
AHQDKMGGMDALHAAGIATYANALSNQLAPQEGHVAQAQSLTFAANGWEPATAPNFGPL
KVFPYGPGHSDNITVGDGTDIAFGGCLIKDSKAKSLGNLGDADTEHYAASARAFGAAF
PKASHIVMSHSAPDSRAAITHTARMADKLR
```



Example – ResFinder

ResFinder 3.0

ResFinder identifies acquired antimicrobial resistance genes and/or find chromosomal mutations in total or partial sequenced isolates of bacteria.

The database is curated by: **Valeria Bortolaia** (click to contact)

View the [version history](#) of this server.

Chromosomal point mutations

Resistance caused by mutations
Select species
Salmonella

Show unknown mutations
Show only known mutations
Show all mutations, known and unknown

Acquired antimicrobial resistance genes

Select Antimicrobial configuration
Select multiple items, with Ctrl-Click (or Cmd-Click on Mac) - by default all databases are selected

- Aminoglycoside
- Beta-lactam
- Colistin
- Fluoroquinolone
- Fosfomycin
- Fusidic Acid

Select threshold for %ID
90 %

Select minimum length
60 %

Draft genome of *S. Typhi* TY_5359 (CIP^S and ESBL producing)

<https://cge.cbs.dtu.dk/services/ResFinder/>



ResFinder-3.0 Server - Results

Input Files: *Strain1.fsa*

[Show Point mutation results](#)

Chromosomal point mutations - Results

Species: *salmonella*

Known Mutations

parE

No mutations found in parE

parC

No known mutations found in parC

gyrA

No known mutations found in gyrA

pmrB

No known mutations found in pmrB

pmrA

No known mutations found in pmrA

gyrB

No mutations found in gyrB

16S_rrsD

No known mutations found in 16S_rrsD

23S

No known mutations found in 23S



BLAST to ResFinder database (assemblies)

Center for Genomic Epidemiology

Username: AMHammeru
Password: *****
New Reset Login

Home Services Instructions Output Overview of genes Article abstract

ResFinder 3.0

ResFinder identifies acquired antimicrobial resistance genes and/or find chromosomal mutations in total or partial sequenced isolates of bacteria.

The database is curated by: **Valeria Bortolaia**
(click to contact)

View the [version history](#) of this server.

Chromosomal point mutations

Acquired antimicrobial resistance genes

Select Antimicrobial configuration
Select multiple items, with Ctrl-Click (or Cmd-Click on Mac) - by default all databases are selected

- Aminoglycoside
- Beta-lactam
- Colistin
- Fluoroquinolone
- Fosfomicin
- Fusidic Acid

Select threshold for %ID: 90 %

Select minimum length: 60 %

Select type of your reads: Assembled Genome/Contigs*

If you get an "Access forbidden. Error 403": Make sure the start of the web address is https and not just http. Fix it by clicking [here](#).

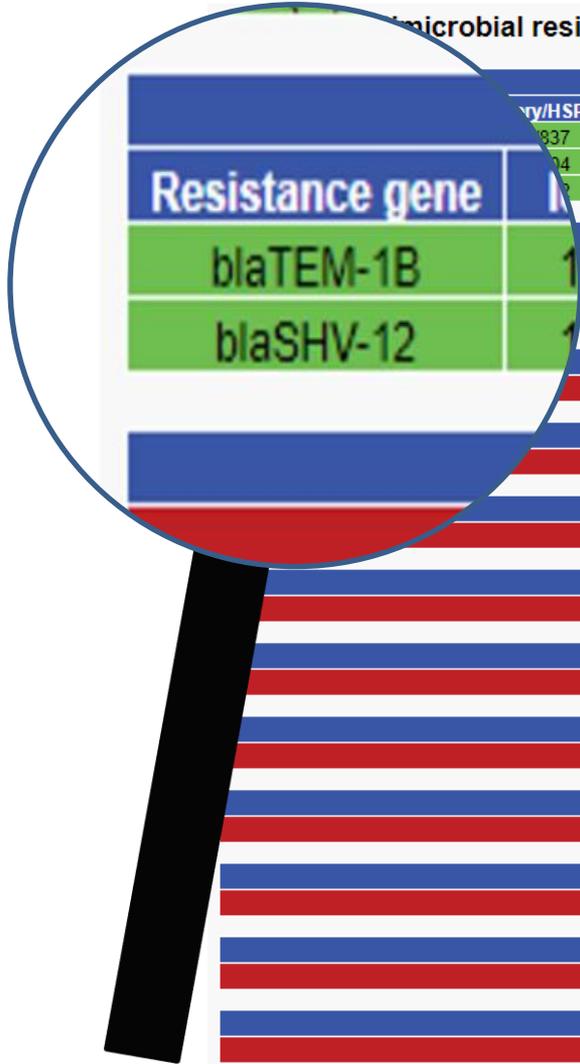
Isolate File

Name	Size	Progress	Status
Strain1.fsa	5.12 MB	<div style="width: 100%;"></div>	

Draft genome of *S. Typhi* TY_5359
(CIP^S and ESBL producing)



Antimicrobial resistance gene - Results



Aminoglycoside						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
		837	S_Typhi_TY_5359_R1_(paired)_contig_62	1635..2471		M28829
		94	S_Typhi_TY_5359_R1_(paired)_contig_62	2471..3274	Aminoglycoside resistance	M96392
			S_Typhi_TY_5359_R1_(paired)_contig_78	278..859	Aminoglycoside resistance	AF162771

Beta-lactam						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
blaTEM-1B	100.00	1185/1185	S_Typhi_TY_5359_R1_(paired)_contig_11	88530..89390	Beta-lactam resistance	JF910132
blaSHV-12	100.00	570/570	S_Typhi_TY_5359_R1_(paired)_contig_46	93..953	Beta-lactam resistance	AF462395

Colistin
No resistance genes found.

Fluoroquinolone
No resistance genes found.

Fosfomycin
No resistance genes found.

Fusidic Acid
No resistance genes found.

Glycopeptide
No resistance genes found.

MLS - Macrolide, Lincosamide and Streptogramin B
No resistance genes found.

Nitroimidazole
No resistance genes found.

Oxazolidinone
No resistance genes found.

Phenicol
No resistance genes found.

Rifampicin
No resistance genes found.

Sulphonamide
No resistance genes found.

Tetracycline						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
tet(D)	100.00	1185/1185	S_Typhi_TY_5359_R1_(paired)_contig_75	721..1905	Tetracycline resistance	AF467077

Trimethoprim						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
dfpA18	99.82	570/570	S_Typhi_TY_5359_R1_(paired)_contig_64	4741..5310	Trimethoprim resistance	AJ310778



UNIX tools – SRST2

Installation

1 - Install dependencies

- python (v2.7.5)
- scipy <http://www.scipy.org/install.html>
- bowtie2 v2.1.0 <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>
- SAMtools v0.1.18 <https://sourceforge.net/projects/samtools/files/samtools/0.1.18/> (NOTE 0.1.19 DOES NOT WORK)

N.B. If you have multiple versions of samtools or bowtie2 installed, you can pick which one `srst2` or `slurm_srst2` should use by setting the following environment variables.

- `SRST2_SAMTOOLS`
- `SRST2_BOWTIE2`
- `SRST2_BOWTIE2_BUILD`

If these aren't set or are missing, they will default to looking in your `PATH` for `samtools`, `bowtie2` and `bowtie2-build`. The exception is `SRST2_BOWTIE2_BUILD` which, if it is not set or missing, will try adding `-build` to `SRST2_BOWTIE2` if it exists, otherwise it defaults to looking in your `PATH`

2 - Get and install the code

Make sure you have installed [git](#) and [pip](#).

Clone the git repository: `git clone https://github.com/katholt/srst2`

and then install with pip: `pip install srst2/`

OR do both at once: `sudo pip install git+https://github.com/katholt/srst2`

3 - Test that the programs are installed properly

```
srst2 --version
getm1st.py -h
slurm_srst2.py -h
```



Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data

Philip T. L. C. Clausen¹, Ea Zankari^{1,2}, Frank M. Aarestrup² and Ole Lund^{1*}

¹Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kgs Lyngby, Denmark;

²Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, 2800 Kgs Lyngby, Denmark

*Corresponding author. Tel: +45 4525 2425; Fax: +45 4593 1585; E-mail: lund@cbs.dtu.dk

Received 26 February 2016; returned 8 April 2016; revised 21 April 2016; accepted 21 April 2016

ResFinder

BLASTn

De novo assembled data
Comprehensive gene db

SRST2

Mapping by Bowtie2
Raw sequencing reads
Reduced db (80% clusters)

KmerResistance (=KMA)

Mapping by kmers
Raw sequencing reads
Mapping to ResFinder db



Materials & Methods

- **Two datasets**
 - Raw NGS data
 - Phenotypic susceptibility test (MIC)
- *Full dataset*
- *Down-sampled dataset*
 - 1% of original reads

DATASET	DK	UK
No. isolates	196	143
Bacterial species	<i>-Salmonella</i> Typhimurium <i>-E. coli</i> <i>-E. faecalis</i> <i>-E. faecium</i>	<i>-E. coli</i> <i>-K. pneumoniae</i>
Host	Pigs	Human



Results

Dataset	Method	SMC*
UK – full set	SRST2	95.92%
	ResFinder	96.04%
	KmerResistance	96.04%
DK – full set	SRST2	97.57%
	ResFinder	97.76%
	KmerResistance	99.61%
UK - downsampled	SRST2	73.43%
	ResFinder	71.33%
	KmerResistance	95.34%
DK - downsampled	SRST2	90.47%
	ResFinder	88.43%
	KmerResistance	99.38%

SMC; simple matching coefficient

* SMC = Simple matching coefficient

Full dataset

Reduced dataset





Prediction of Phenotypic Antimicrobial Resistance Profiles From Whole Genome Sequences of Non-typhoidal *Salmonella enterica*

Saskia Neuert^{1,2}, Satheesh Nair², Martin R. Day², Michel Doumith², Philip M. Ashton², Kate C. Mellor^{3,4}, Claire Jenkins^{1,2}, Katie L. Hopkins², Neil Woodford², Elizabeth de Pinna², Gauri Godbole^{1,2} and Timothy J. Dallman^{1,2*}

- 3491 Non-typhoidal isolates (227 serovars)
- Availability of both phenotypic data and NGS data
- Fifteen antimicrobials including ciprofloxacin
- Mapping (Bowtie 2) of raw reads to homemade database



TABLE 1 | Comparison of phenotypic antimicrobial susceptibility testing and genome-derived resistance prediction for non-typhoidal *Salmonella enterica* (*n* = 3491).

Antimicrobial	Phenotype: susceptible		Phenotype: resistant		Sensitivity (%)	Specificity (%)
	Genotype: resistant	Genotype: susceptible	Genotype: resistant	Genotype: susceptible		
Ampicillin	1	2742	747	1	99.87	99.96
Temocillin	0	3490	0	1	0	100
Cefoxitin	0	3471	19	1	95.0	100
Cefotaxime	0	3434	57	0	100	100
Ceftazidime	0	3444	47	0	100	100
Cefpirome	0	3444	47	0	100	100
Ertapenem	0	3481	10	0	100	100
Chloramphenicol	4	3284	201	2	99.01	99.88
Gentamicin	1	3351	138	1	99.28	99.97
Streptomycin	51	2821	613	8	98.71	98.22
Tobramycin	2	3392	97	0	100	99.94
Sulphonamides	2	2661	828	0	100	99.92
Tetracycline	6	2568	917	0	100	99.77
Trimethoprim	1	3185	301	4	98.69	99.97
Ciprofloxacin	1	3352	137	1	99.28	99.97

Values shown designate the number of isolates. For ciprofloxacin only isolates with a MIC > 0.5 mg/L are shown.

Mainly gene based resistance (transferable)

Both mutation based (chromosomal) and gene based resistance (transferable)





ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

International Journal of Antimicrobial Agents

journal homepage: <http://www.elsevier.com/locate/ijantimicag>



Letters to the Editor

Use of whole-genome sequencing for characterisation of a ST119 NDM-1-producing *Acinetobacter pittii* from a patient in Denmark with no history of recent travel

Keywords:

Carbapenemase
Metallo-beta-lactamases
Multiresistant



Anette M. Hammerum*

Frank Hansen

Department of Microbiology and Infection Control, Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen S, Denmark

Pia Littauer

Department of Clinical Microbiology, Hvidovre University Hospital, Hvidovre, Denmark

* Corresponding author. Tel.: +45 3268 3399; fax: +45 3268 3231.

E-mail address: ama@ssi.dk (A.M. Hammerum)

29 May 2015

<http://dx.doi.org/10.1016/j.ijantimicag.2015.05.005>

In December 2014, a meropenem-resistant *Acinetobacter* spp. was isolated from the abdominal tissue from a 73-year-old male patient at the Department of Clinical Microbiology at Hvidovre University Hospital (Hvidovre, Denmark). The isolate was identified as *A. pittii* by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF).

The *bla*_{NDM} gene was identified by PCR from the *A. pittii* isolate.



	MIC
Betalactam	
IMI	>8
A/S	>16
CAR	>256
TIC	>64
TIM	> 128
PIP	>64
P/T	>64
FOT	>32
AXO	>32
TAZ	>16
FEP	>16
FOP	>32
AZT	>16
Aminoglycoside	
GEN	>8
AMI	≤ 4
TOB	≤ 1
Fluoroquinolone	
e	
CIP	>2
LOM	>4
LEVO	>4
Sulphonamide	
Trimethoprim	
SXT	>4
FIS	>256
Tetracycline	
TET	>8
Phenicol	
CHL	>16

ResFinder-3.0 Server - Results

Input Files: *CPO20140051*

[Show Acquired antimicrobial resistance results](#)

Acquired antimicrobial resistance gene - Results

Aminoglycoside						
No resistance genes found.						
Beta-lactam						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
blaNDM-1	100.00	813/813	CPO-2014-0051_S9_L001_R1_001_(paired)_contig_54	340..1152	Beta-lactam resistance	FN398879
blaADC-25	91.84	1152/1152	CPO-2014-0051_S9_L001_R1_001_(paired)_contig_18	98105..99256	Beta-lactam resistance	EF016355
Colistin						
No resistance genes found.						
Fluoroquinolone						
No resistance genes found.						
Fosfomycin						
No resistance genes found.						
Fusidic Acid						
No resistance genes found.						
Glycopeptide						
No resistance genes found.						
MLS - Macrolide, Lincosamide and Streptogramin B						
No resistance genes found.						
Nitroimidazole						
No resistance genes found.						
Oxazolidinone						
No resistance genes found.						
Phenicol						
No resistance genes found.						
Rifampicin						
No resistance genes found.						
Sulphonamide						
No resistance genes found.						
Tetracycline						
No resistance genes found.						
Trimethoprim						
No resistance genes found.						

Time is life (and money)

Isolation
Purification
Sequencing

4-7 days



Relatively..

- low price per isolate
 - well-proven technology
 - high precision (low error rate)
 - **slow**
- ..and no reads in real-time**



6-24 hours

Relatively..

- **expensive**
 - **experimental technology**
 - **low precision (high error rate)**
 - fast
- ..and reads available in real-time**



Illumina vs ONT MinION data

Genome



Illumina reads
(Short)
(low error rate)



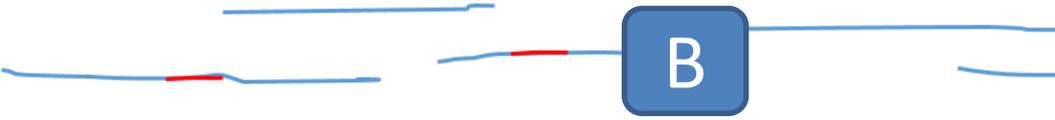
Illumina assembly



Error rate 0.1 – 1 %

Error rate < 0.001%

MinION reads
(Long)
(high error rate)



Error rate 5 – 15 %

Error rate 5 – 10 %

MinION assembly



— Repeat area (rRNA, IS, homologue genes ect..)



A: Illumina assembly

ResFinder-3.0 Server - Results

Input Files: *CPO20160062.fa*

Show Acquired antimicrobial resistance results

Acquired antimicrobial resistance gene - Results

Aminoglycoside						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
aac(6')Ib-cr	100.00	519/519	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_15_cons	146..664	Fluoroquinolone and aminoglycoside resistance	EF636461
aph(3)-Ia	100.00	816/816	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_124_con	155..970	Aminoglycoside resistance	V00359
aph(6)-Id	100.00	837/694	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_92_consen	1772..2465		M28829
aac(3)-IIa	100.00	861/861	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_144_con	64..924	Aminoglycoside resistance	CP023555
			CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_92_consen	969..1772		AF321551

Beta-lactam						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
blaNDM-1	100.00	813/813	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_139_conse	243..1055	Beta-lactam resistance	FN396876
blaOXA-1	100.00	831/831	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_15_consen	795..1625	Beta-lactam resistance	J02967
blaLEN12	91.21	789/705	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_138_con	79..783	Beta-lactam resistance	AJ635406
blaTEM-1B	100.00	861/861	CPO20160062_S1_L001_R1_001_(paired)_trimmed_(paired)_contig_67_consen	873..1733	Beta-lactam resistance	JF910132



B: MinION reads

Acquired antimicrobial resistance gene - Results

Beta-lactam						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
blaTEM-1B	90.53	861/876	dade357f-7ca2-4b86-9454-c5a356274fd0	10026..10867	Beta-lactam resistance	JF910132
blaNDM-1	93.15	813/818	757f96b2-06b0-4a8e-aa51-2c526f39f12b	100608..101408	Beta-lactam resistance	FN396876
blaSHV-11	91.92	861/879	48e1de69-20ec-4ab4-9caf-74e4aee8e01e	10162..11018	Beta-lactam resistance	GQ407109
blaSHV-28	91.37	831/845	1ba9cf7d-7ffe-4908-a524-831aa532da51	10315..11134	Beta-lactam resistance	J02967
blaOXA-1	91.73	861/874	f3ad7fe0-50f4-4eeb-89d3-e0a2b5e4f1be	10330..11183	Beta-lactam resistance	JF910132
blaNDM-6	92.82	813/822	a38d41ff-f86d-482f-bde1-21b0286b2305	1046..1898	Beta-lactam resistance	HM751101
blaNDM-11	90.02	813/822	c4ca51de-2ef3-41c1-a0b2-593371313bb2	10573..11302	Beta-lactam resistance	J02967
blaNDM-1	90.25	813/822	e5c685a1-c4f5-45bb-aa0e-7d86f4596563	1065..1867	Beta-lactam resistance	JN967644
blaOXA-1	93.15	831/822	f73c3c2b-8c12-4d7a-ad36-f74e09a003c9	10757..11561	Beta-lactam resistance	KP265939
blaOXA-1	92.83	831/822	b12f5a9e-0930-4ac4-8b24-3dff5a8e6044	10901..11682	Beta-lactam resistance	FN396876
blaOXA-1	92.83	831/822	32089238-879d-4407-911c-107c2cc17971	11032..11829	Beta-lactam resistance	J02967
blaOXA-1	92.83	831/822	0252c202-91d7-4d26-a1ea-a178e80881cc	11083..11888	Beta-lactam resistance	J02967
blaOXA-1	92.83	831/822	2ea7e5f3-0f3a-416a-9d88-110fc6590f0b	11093..11888	Beta-lactam resistance	DQ286729
blaOXA-1	92.83	831/822	04b7ee1a-9b41-4e2f-87e9-b29f4a919ae4	11168..12016	Beta-lactam resistance	JF910132
blaOXA-1	92.83	831/822	77ddc41d-4e84-4515-b48e-df2de99c6ca0	11179..12007	Beta-lactam resistance	AF462395
blaOXA-1	92.83	831/822	bcd10027-7f8f-472a-a26e-1aa9b066888d	11198..12043	Beta-lactam resistance	AF190694
blaOXA-1	92.83	831/822	7d099581-6767-49ce-bc8b-ef6c7fa23cff	11232..12075	Beta-lactam resistance	AF188199
blaTEM-55	91.72	831/845	08288484-896b-4698-9306-336aa08e9b5e	11388..12202	Beta-lactam resistance	J02967
blaTEM-1B	90.53	831/850	880b828d-a4ba-4ae7-b8d9-9b59e6e37b5c	11510..12342	Beta-lactam resistance	J02967
blaTEM-1B	90.53	861/848	c3d9798e-8ecc-46bb-a113-5c0cc318e6f6	11645..12453	Beta-lactam resistance	JF910132
blaTEM-198	92.47	861/877	78a58875-c54a-4aa2-ae3f-c50294e9a9ef	11687..12539	Beta-lactam resistance	AB700703
blaOXA-1	91.10	831/831	5467ded8-2bd6-4669-aad2-99cb7614a23c	11827..12626	Beta-lactam resistance	J02967
blaOXA-1	90.35	831/850	4a657114-aa94-4b01-bf56-7c0acccd1e5e	11837..12659	Beta-lactam resistance	J02967
blaNDM-1	92.29	813/817	c95a826e-eeb5-4639-8e95-f3d0118be2c6	11851..12649	Beta-lactam resistance	FN396876

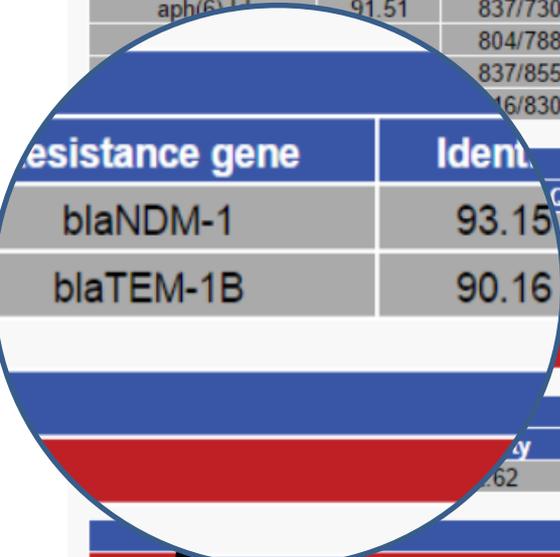
Error rate around 5-15%....impossible to identify specific genes.



C: MinION assembly

Acquired antimicrobial resistance gene - Results

Aminoglycoside						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
aph(6)-Id	90.96	837/697	2	183990..184662		M28829
aph(3")-Ib	90.70	804/828	2	184666..185471		AF321551
aac(6')Ib-cr	92.62	519/488	2	34323..34788	Fluoroquinolone and aminoglycoside resistance	EF636461
aph(6)-Id	91.51	837/730	2	52359..53066		M28829
		804/788	2	82180..82940		M28829
		837/855	2	82948..83772		M28829
		16/830	2	84675..85475	Aminoglycoside resistance	V00359



Resistance gene	Identity	Beta-lactam				
		Query/HSP	Contig	Position in contig	Phenotype	Accession no.
blaNDM-1	93.15	813/818	3	163028..163828	Beta-lactam resistance	FN396876
blaTEM-1B	90.16	861/884	2	181022..181874	Beta-lactam resistance	JF910132

Colistin
No resistance genes found.

Fluoroquinolone						
Resistance gene	Identity	Query/HSP	Contig	Position in contig	Phenotype	Accession no.
	92.62	519/488	2	34323..34788	Fluoroquinolone and aminoglycoside resistance	EF636461

Fosfomycin
No resistance genes found.

Fusidic Acid
No resistance genes found.

Error rate around 5-10%....impossible to identify specific genes.



SCIENTIFIC REPORTS

OPEN

Developing an *in silico* minimum inhibitory concentration panel test for *Klebsiella pneumoniae*

Received: 27 September 2017

Accepted: 12 December 2017

Published online: 11 January 2018

Marcus Nguyen^{1,2,3}, Thomas Brettin^{2,3}, S. Wesley Long^{4,5}, James M. Musser^{4,5}, Randall J. Olsen^{4,5}, Robert Olson^{2,3}, Maulik Shukla^{2,3}, Rick L. Stevens^{2,3,6}, Fangfang Xia^{2,3}, Hyunseung Yoo^{2,3} & James J. Davis^{2,3}

- 1668 clinical isolates
- Availability of both MIC data and NGS data
- Twenty antimicrobials including ciprofloxacin
- Kmer-based machine learning → *no a priori* knowledge about genes

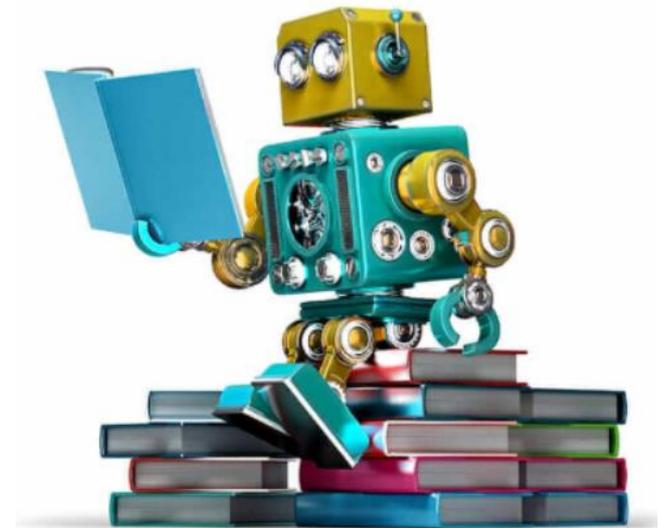


Table 1. Accuracies for the entire XGBoost model and for the individual antibiotics.

Antibiotic	Samples	Accuracy ^a	95% C.I. ^b
All	32705	0.92	0.92, 0.92
Amikacin	1667	0.97	0.96, 0.98
Ampicillin	1666	1.00	0.99, 1.00
Ampicillin/Sulbactam	1664	0.99	0.99, 1.00
Aztreonam	1644	0.89	0.89, 0.90
Cefazolin	1667	0.96	0.95, 0.96
Cefepime	1571	0.61	0.58, 0.64
Cefoxitin	1645	0.90	0.89, 0.91
Ceftazidime	1667	0.92	0.91, 0.93
Ceftriaxone	1667	0.89	0.87, 0.90
Cefuroxime sodium	1575	0.99	0.99, 1.00
Ciprofloxacin	1664	0.98	0.97, 0.98
Gentamicin	1667	0.95	0.93, 0.96
Imipenem	1666	0.94	0.93, 0.95
Levofloxacin	1666	0.97	0.96, 0.97
Meropenem	1660	0.93	0.91, 0.95
Nitrofurantoin	895	0.96	0.95, 0.97
Piperacillin/Tazobactam	1662	0.78	0.77, 0.79
Tetracycline	1667	0.89	0.87, 0.90
Tobramycin	1666	0.95	0.94, 0.96
Trimethoprim/Sulfamethoxazole	1667	0.95	0.94, 0.96

^a Average accuracy within ± 1 two-fold dilution factor, based on a 10-fold cross validation.

^b 95% confidence interval.



Conclusions

NGS is a wonderful tool to study AMR, but...

- “Known knowns” are easier to predict than “Known unknowns”
- Define your need first...then find the best tool(s) to get there
- Read the paper(s) describing the tool(s) you select
- Consider to hire a bioinformatician

Thank you for your
attention

Questions?

